Standard Middleware Services

- "Refine the definition of a Grid Services Package designed to be deployed across testbeds and used by applications"
  - Ruth Aydt
  - Ian Foster
  - Bill Johnston
  - Miron Livny
  - Jo Sander
  - Rick Stevens
  - Mary Thompson

The Basic Idea (EMERGE Proposal)

- Deploy standard infrastructure across sites
  - Provide maximum capabilities to applns
  - Increase what can be “taken for granted” when developing applications
  - Reduce deployment burden at sites
- For example
  - Authentication, resource discovery, resource management, instrumentation, ….
- Call this a “Grid Services Package”
- Persistence is the key

Observations on EMERGE

- EMERGE is DOE NGI funded but is multi-institutional
  - NSF PACI, Universities, I-2, DOE labs
- EMERGE integrates technologies developed under funding from multiple sources
  - NSF, NASA IPG, DARPA, DOE (e.g., Clipper)
  - Tightly coordinated due to bottom-up organization of Grid community
- DOE NGI funding allows extensions to address QoS and other issues

Two Fundamental Services

- Authentication infrastructure
  - Certificate authority(s) for identity
  - Support for 3rd-party service for sites that do not run their own CAs
  - Estimate 1-2 FTE effort to run this in ESnet
- Information infrastructure
  - State information and name service
  - Scalable, reliable implementation
  - Support for 3rd-party service
  - Estimate 3+ (?) incremental FTEs

Other Services

- Instrumentation infrastructure
  - Integrated with information service
- High-speed data transport service
  - Transparent access to low-level mechanisms: striping, sack, parameter tuning, etc.
- Resource management service
  - Implementations and APIs for QoS delivery
  - Resource tracking and accounting
- Other services from DOE NGI projects and elsewhere

Approach

- Use EMERGE as a focal point for early development, deployment, evaluation, demonstration (Year 1)
- Work to create authentication, information infrastructure for other NGI sites (Year 1?)
- Plan to deploy Grid Services Package to other sites, if proven successful in EMERGE (Year 2)

Other Slides

Other Issues

- Maintaining a distributed system
- Code updates, system monitoring, etc.
  - Centralized “push” of new software
  - Protocol numbers in key components
  - “Heartbeat monitors”
- End-to-end application-level encryption

New Things that are Needed

- CA needed for issuing identity certificates for NGI participants outside EMERGE
- Accounting infrastructure for network
  - “Grid system logs” (e.g., resource manager requests) accessible via the uniform publication and access methods being defined by instrumentation group
  - Also summaries on per-user, per-project, etc., usage
  - Resource managers need to generate this information

Other Requirements

- High-speed transport as a middleware svc
  - Via striping, parameter tuning, sack, etc.
  - Or maybe specialized protocols
  - Transparent to the user
  - Can’t throw away congestion control!
- QoS services
  - Guaranteed delivery of bulk data
  - Advance reservation of premium bandwidth
  - Advance reservation of low-latency paths

Questions

- What middleware services are required for DOE NGI applications? What are priorities?
- What middleware services are available to us now? What will become available over the course of the project from DOE NGI activities?
- Are there key middleware services that are not available?
- What are the obstacles to the use of middleware services?
- How can these be overcome?
- What do middleware people need from testbeds?
- What do testbeds want from middleware people?
- Should we be talking about a coordinated approach to the deployment of middleware services? (as EMERGE is doing with its Grid Services Package). If so, who should handle this?

Basic Services

- Identity based on global credential
- Authentication & authorization at resource managers via GSS-API and GAA
- Accounting of premium network usage
  - Per user tied to global identity
  - Network usage on a per-project basis
  - Characterize usage of network
  - Quantify successful and unsuccessful requests